

# OCEAN RESEARCH SYNTHESIS AND MODELING PROGRAM (ORSMP): A PROGRESS REPORT.

Thomas M. Powell

30 October 1998

## I. INTRODUCTION

In the last eighteen months community meetings "...have addressed the proposition that the present state of ocean modeling and data assimilation is profoundly unsatisfactory." The largest and most recent of these was held at the National Science Foundation, Washington, D.C., 4-5 May 1998. Meeting reports from two earlier gatherings (U.S. WOCE Office 1997a, b) are available summarizing their findings, as well as a published article (Nowlin 1997). A discussion document (Powell et al. 1998) for the May meeting is attached as Appendix B.

Nowlin (1997) detailed the attempts to initiate an appropriate response to this important and identifiable need. Powell et al. (1998) concluded that:

"In short, arguments to begin a substantial enhancement of modeling and data assimilation capabilities in ALL sub-disciplines of the ocean sciences are compelling. The most critical of these reasons involve: existing and new satellites and the data they are (will be) collecting; the massive data sets assembled by WOCE, JGOFS, GLOBEC, etc.; the needs for coordination among modelers, and between modelers and observationalists; and the requirement for greater, yet more diverse, computing capability. These facts are well-documented (U.S. WOCE Office 1997a, U.S. WOCE Office 1997b) and are not in dispute."

In addition, as sketched in Nowlin (1997) and described by Powell et al. (1998), there has emerged a broad community consensus on the structure that will best address this need: a central facility, termed "the hub", linked closely to a series of research endeavors of varying size and complexity, termed "nodes" in the earlier reports. Working groups at the May 1998 meeting reviewed the "hub-node" architecture. They added a number of important details to the notions of each, but retained the basic "hub-node" structure. In addition, one working group considered the possible organization and infrastructure (including outreach and educational functions) that such a structure will demand.

Building on the conclusions from the earlier meetings and documents, this brief interim progress report summarizes the conclusions that emerged from the May meeting. This short report ends with five recommendations for steps needed to initiate this program.

## II. THE CENTRAL "HUB" FACILITY.

Very brief descriptions of the technical, computational, and data capabilities thought to be needed for the central hub were suggested by Nowlin (1997) and Powell et al. (1998). Participants at the May 1998 meeting re-confirmed these basic constituents:

- 1) computing cycles
- 2) data streams, including

- 3) high-level analyses,
- 4) some real-time acquisition
- 5) technical assistance
- 6) code and analysis software
- 7) benchmark solutions
- 8) directories and documentation

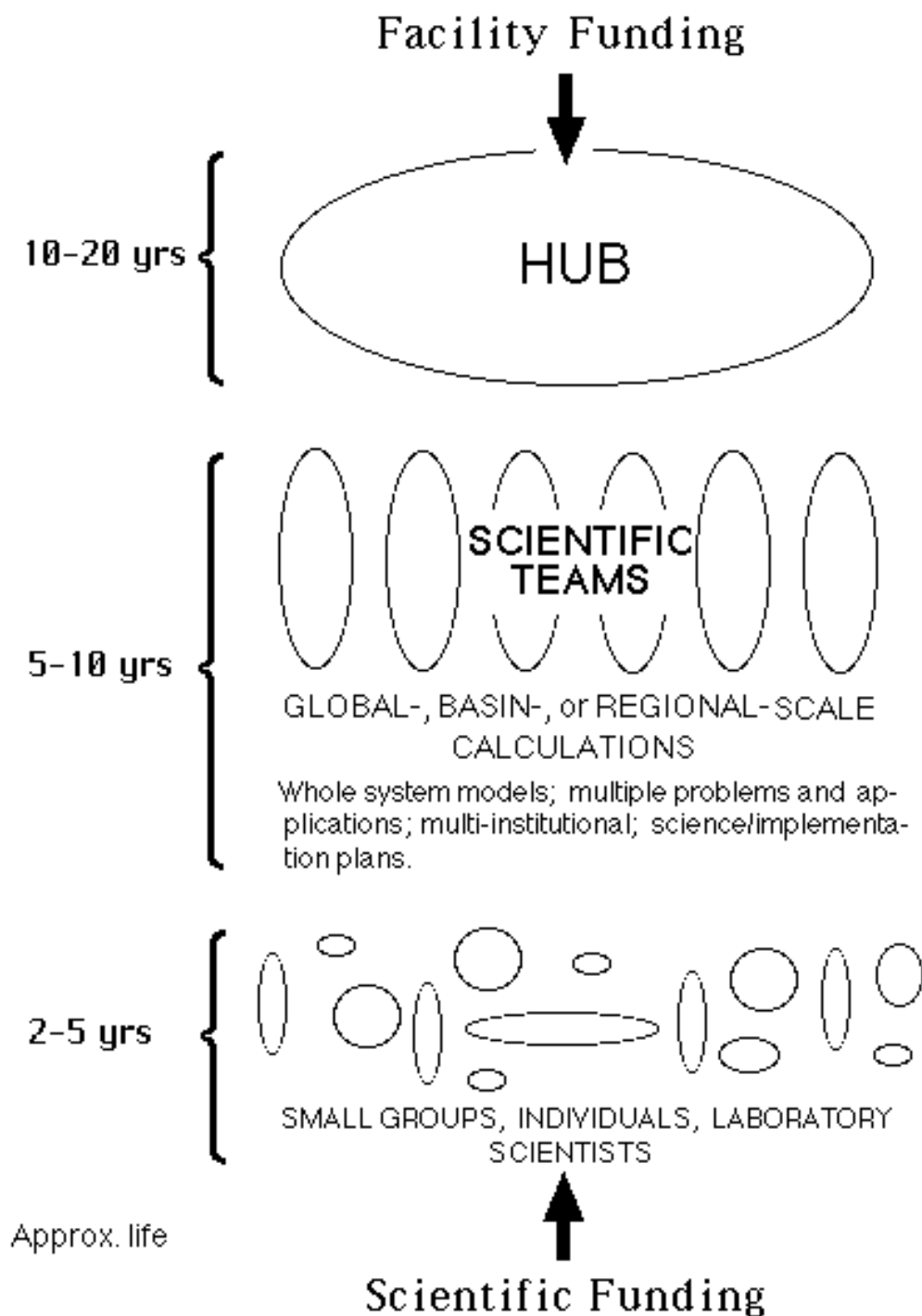
An optimal location for the hub would be at a neutral site; that is, not co-located with a large, dominant modeling or data assimilation group. Moreover, pro-active scientific oversight of the hub is a critical prerequisite to effective functioning of the hub...and the entire hub-node complex. Certain activities should initially take place outside the hub. These include model improvement and simulations, and (data assimilative) estimates of the state of the ocean. But it should be a goal of the program to bring these capabilities into the hub (and to operational agencies, as appropriate) once they mature.

Consideration of the hub's relation to the nodes led participants to consider an amplification of the simple hub-node structure of Nowlin (1997) and Powell et al. (1998). The complexity and duration of some scientific projects may demand nodes of quite large size. It may be useful to classify large scientific teams separately from the activities of small groups, or individual scientists. Figure 1 depicts such a separation, including suggested activities that might distinguish scientific teams from the smaller groups. Figure 1 also depicts two funding streams: first, a facility funding avenue that must be the source of support for the hub; and second, the traditional scientific funding routes to which small groups and individuals apply. Though the large scientific teams would be expected to obtain the bulk of their funds through the usual scientific funding modes, perhaps some activities can be supported via facility funding, perhaps in conjunction with the hub. [N.b., participants strongly endorsed the notion that even large scientific teams must remain substantially smaller than existing major ocean programs (e.g., WOCE, JGOFS, GLOBEC)].

An important final message of figure 1 is that both scientific teams and small groups are "nodes". And any hub linked to the nodes must be designed to accommodate both categories of users -- large and small.

### **III. NODES.**

Powell et al's (1998) discussion document was largely devoted to the nature of the node concept, giving four detailed examples of possible nodes. These were: 1) the coastal ocean; 2) coupled physical-biological models; 3) marine geochemistry and biogeochemistry; 4) general circulation of the ocean. Position papers on these four potential nodes were presented in plenary session to the May meeting participants. See Appendix A. Following the classification of Figure 1, these four would qualify as scientific teams. Powell et al. also argued that the entire program should be "node-driven science", concluding their document with:



**Figure 1.** The central hub facility and nodes of various sizes, lifetimes.

"Though the most constructive scientific discourse always involves give-and-take -- in this case between "hub" and nodes -- requirements from the nodes should shape the "hub", and not the converse."

Participants at the May 1998 meeting strongly supported this posture. The hub was seen as a facilitator of the nodes; it was to be expected that the nodes would generate the ideas and themes that the overall program should pursue. One might not be surprised to find that communication between nodes might be far greater than between the hub and the node.

Several (desirable) elements that might characterize a node were suggested.

- Defined by a research problem (needing hub services).
- Ca. 5 year lifetime.
- Large in scope.
- Multi-disciplinary activities encouraged.
- Open access over time (i.e., new PIs)
- \*\* GOOD SCIENCE!
- \*\* BUILDS COMMUNITY MODELING CAPABILITY.

The last two elements (marked by double asterisks) were considered essential for an activity to be included as a node within the overall program.

A number of specific research problems that will demand major modeling and/or data assimilation efforts -- i.e., nodes -- were suggested. They appear in Figure 2. [The examples of nodes/scientific teams from Powell et al. (1998) are also listed in Figure 2.] Participants speculated that the research problems shown in Figure 2 would need groups of 5-15 PIs to insure adequate progress. One should acknowledge that the boundary between the "scientific teams" and the "small groups" is not a rigid one. Perhaps 5 PIs qualifies as small group; perhaps 5 PIs comprise a scientific team, depending upon the nature of their project. It seems wise to retain this flexibility, if possible.

Several questions about the nature of the nodes arose in discussion. Should a data assimilation component be a requirement for a node project? Most thought not. A close connection to the observational community is desirable. Should it be required? Should "management" insist that node projects develop a relationship with the operational community? Finally, should there be a proscribed "life-history" for a node? A fixed lifetime? Should a "sunset clause" be enforced to insure that new talent enters the program? Again flexibility seems warranted at this early stage. But a future Scientific Advisory Committee can expect to grapple with these questions.

#### IV. ORGANIZATION, INFRASTRUCTURE.

One working group at the May 1998 meeting considered organizational needs for the proposed hub-node system. The group suggested a name: Ocean Research Synthesis and Modeling Program (ORSMP). They also submitted a statement of purpose for ORSMP which appears below.

##### ORSMP GOALS

- To foster and support the variety of model/data SYNTHESIS projects needed to advance our capability in three endeavors:
  1. Simulate and understand the physical, chemical, biological, and biogeochemical behavior of the ocean.
  2. Estimate the state of the ocean.
  3. Propose new observing activities.
- To promote the development of communities with interests and skills appropriate to accomplish ORSMP goals.
- To provide the infrastructure (computing, networks, data sets, interactive software) needed to support the ORSMP PROJECTS and to remain state-of-the-art.

Speakers from the working group explained that the proposed second goal (“...development of communities...”) was designed to include activities like visiting scientist programs, tutorials, and other outreach and educational programs. Further, the proposed third goal (“...infrastructure...”) reaches beyond a state-of-the-art hub facility. It extends to the nodes, the ORSMP PROJECTS, through its call for state-of-the-art support of these efforts. Most important, however, is the phrase “...model/data SYNTHESIS projects...” in the proposed first goal. Participants placed the highest priority on synthetic endeavors, and argued that this was also the sense of the broader community. Though all acknowledge the importance of the development of models and assimilative schemes, to do so in a data-free vacuum should not be the sole aim of ORSMP projects.

The working group also weighed in with six specific recommendations. FIRST, NSF should be the lead agency for ORSMP. But participants placed a high priority on the development of interagency coordination, and, hopefully, interagency sponsorship. SECOND, NSF should appoint a Science Advisory Group immediately, which should meet regularly in open meetings and report on progress to the larger community. THIRD, the Advisory Group should search for an ORSMP leader. Once named, the leader should develop a plan for the ORSMP hub facility in conjunction with the NSF, other agencies (if jointly desired), the Advisory Group, and appropriate members of the broader community. FOURTH, the hub facility should be “lean”, especially in numbers of scientific staff. Moreover, of the several hub characteristics listed on pp. 1, 2, this working group singled out computing cycles, data streams, and interactive software for special attention. They also made a plea for a systematic upgrade process within hub functions. FIFTH, substantial attention should be directed early to outreach efforts of all kinds...at both the hub and the nodes. SIXTH, the Science Advisory Group should develop a strategy to solicit and review ORSMP PROJECT proposals. This working group felt the synthetic nature of desirable ORSMP PROJECTS – data/model SYNTHESIS – might present both difficulties

## NODES. SCIENTIFIC TEAMS

The coastal ocean

Coupled physical-biological models

Marine geochemistry and biogeochemistry

Ocean general circulation

## NODES. RESEARCH PROBLEMS

NAO: climate, fisheries

Coastal circulation  $\longleftrightarrow$  larval transport  
(site-specific; compare several sites)

Decadal variability of the ENSO phenomenon

Global  $\left\{ \begin{array}{l} \text{C} \\ \text{N} \\ \text{P} \end{array} \right.$  sources/sinks; inverse modeling

Glacial/interglacial transition

Ice-ocean modeling

**Figure 2.** Possible nodes; research problems that investigators at these (and other) nodes might address.

and unique opportunities. Those proposals that could contribute greatly to ORSMP goals because of their distinct integration of data and modeling themes might be reviewed poorly.

## **V. CONCLUDING REMARKS. RECOMMENDATIONS.**

The invitation to the May 1998 meeting set out two goals for the attendees. FIRST, review and revise existing documents and, in the process, propose a plan of action to meet the goals for a proposed Ocean Research Synthesis and Modeling Program. SECOND, recommend minimal steps that must be taken immediately to reach the targets set out in the action plan.

At the conclusion of the meeting those attending agreed without dissent that the basic architecture that had been proposed and discussed was sound. Moreover, the skeleton of both hub and nodes had been sufficiently well-defined that the large task of fleshing out the structure could, and should proceed.

A small number of recommendations, summarized in Figure 3, were advanced as immediate next steps, several of which mirror those proposed by the working group on organization and infrastructure.

### **RECOMMENDATION 1. Appoint a Scientific Steering Committee (Science Advisory Group) for ORSMP.**

There are a number of initial aims the steering group might pursue. One should be to construct detailed plans for the central hub. These plans must be consistent with what is certain to be a peer-reviewed competition for the facility. Another is to develop an Implementation Plan and (long-term) Schedule for the full hub-node system. Several other tasks are connected to the recommendations that follow.

### **RECOMMENDATION 2. Form an interagency task team to work informally with the Steering Committee.**

Liaison with the broad suite of federal agencies that support oceanographic research is an important goal for this effort. Co-ordination of modeling and data assimilation activities at this early period is a critical requirement for a successful ORSMP endeavor.

### **RECOMMENDATION 3. Set the procedures in motion to generate an Announcement of Opportunity for two or more nodes.**

The peer review proposal evaluation process will ultimately select those node proposals which will be funded. Nonetheless, it would be valuable to have at least one “physical oceanography” node, because the discipline is the most advanced in modeling and data assimilation. Further, at least one initial node, hopefully more, should address a multi-disciplinary problem. It is important at this early stage to involve researchers from several disciplines. ORSMP must be a program that includes ALL specializations in oceanography, not just the leading modelers and/or data assimilation practitioners. It would also be useful to support a node that is allied to existing (or proposed) programs that involve modeling and/or data assimilation. One candidate might be a scientific team/node devoted to the general circulation of the ocean (see Powell et al. 1998) with a possible connection to GODAE (the Global Ocean Data Assimilation Experiment – see Nowlin 1997).

## RECOMMENDATIONS

- Appoint a Scientific Steering Committee (Science Advisory group) for ORSMP.
- Form an interagency task team to work informally with the Steering Committee.
- Set the procedures in motion to generate an Announcement of Opportunity for two or more nodes.
- Form a working group, or subcommittee of the Scientific Steering Committee, to address the topic of data bases/data archiving for data assimilation.
- Prepare an attractive, but authoritative and concise, description of the ORSMP effort.

**Figure 3.** Recommendations from a community workshop held at NSF, 4-5 May 1998, on ocean modeling and data assimilation/synthesis capability.

**RECOMMENDATION 4. Form a working group, or subcommittee of the Scientific Steering Committee, to address the topic of data bases/data archiving for data assimilation.**

Little attention has been paid to this subject in the previous meetings and documents. The larger topic of data storage is too broad (and probably too expensive) to tackle in an ORSMP setting. But looking at the subject as a critical element of any data assimilation and model/data synthesis may restrict the investigation to manageable size.

**RECOMMENDATION 5. Prepare an attractive, but authoritative and concise, description of the ORSMP effort.**

Highlight goals, plans and possible activities. Thoughtfully designed works of this kind with compelling graphics are useful on many levels – students, colleagues, scientific managers, the general public. Moreover, the media trumpets about entering the “information age”. And oceanography is a discipline the public finds inherently interesting, adventurous, and romantic. Surely, then, oceanographers can express the excitement they find in what is truly a program for the information age.

**VI. ACKNOWLEDGEMENTS.**

I thank Kristin Bishop for her help in preparing this document.

**VII. REFERENCES.**

- Nowlin, W. D. 1997. U.S. Ocean science needs for modeling and data synthesis: status of a community assessment. *Oceanography*. 10: 135-140.
- Powell, T.M., Nowlin, W.D., Doney, S.C. 1998. A distributed modeling and data assimilation capability for the ocean sciences. A discussion document for Model and Data Assimilation/Synthesis workshop. National Science Foundation, Washington, D.C. 4-5 May 1998. [Appendix B of this report.]
- U.S. WOCE Office. 1997a. Assessing Ocean Modeling and Data Assimilation Requirements. 1. Report on workshop to discuss needs for an ocean data assimilation center.
- U.S. WOCE Office. 1997b. Assessing Ocean Modeling and Data Assimilation Requirements. 2. Report on workshop to discuss community needs for ocean global circulation modeling.

## **APPENDIX A:**

### **Invitation Letter and Agenda for May '98 Workshop on Ocean Modeling and Data Assimilation**

**SUBJECT:** Invitation to an ocean modeling and data assimilation/synthesis workshop

This is to invite you to a community workshop to consider an ocean modeling and data assimilation/synthesis capability. This meeting will be held 4-5 May 1998, at the National Science Foundation in Arlington, Virginia; a tentative agenda is enclosed.

This is part of a continuing activity sponsored by the NSF Division of Ocean Sciences to assess ocean community needs for modeling and data assimilation/synthesis capability. Two community workshops focused on assimilation and modeling were held in 1997. Copies of the resulting reports are enclosed unless you previously received a copy. A short meeting was held in September 1997 to begin planning for this workshop. Enclosed is a copy of an article, soon to appear in The Oceanography Society Magazine, summarizing this activity to date.

A working paper suggesting community needs and actions will be prepared and distributed prior to the meeting. A principal goal of the May workshop is to review and revise this paper and recommend a plan of action. Another goal is to recommend the minimal steps that must be taken now to reach the goals of the action plan.

This workshop will be chaired by Thomas (Zack) Powell. Funding for the meeting will be provided by the Ocean Sciences Division. Logistics will be arranged by Maureen Reap of the U.S. WOCE Office. Please contact her by March 31<sup>st</sup> at , telephone 409/845-1443, or fax 409/845-0888, to accept or decline this invitation.

Support is available through the US WOCE Office for those unable to fund travel to this meeting. Please contact Ms. Reap for information on support. She will make all lodging reservations at the Arlington Hilton Hotel regardless of source of support.

We sincerely hope you will be attending this meeting. We believe there are unmet community needs in the synthesis of data using models and data assimilation, and we believe the results of this meeting will help shape the approach to meeting those needs.

## AGENDA

Workshop on Ocean Modeling and Data Assimilation  
4-5 May 1998  
National Science Foundation  
Arlington, Virginia

### Monday, May 4<sup>th</sup>

- 8:30 Welcome. Introduction of attendees. Adoption of agenda. T. Powell
- 8:40 Background and organization. Purpose of meeting. T. Powell/W. Nowlin
- 9:00 The future of modeling and data assimilation in the ocean sciences: the NSF view. M. Purdy
- 9:20 INVITED PAPERS in PLENARY  
(25-minute presentations, with five additional minutes for questions.)
- “Marine geochemistry and biogeochemistry” Scott Doney
- “The coastal ocean” Dale Haidvogel
- 10:20 BREAK
- 10:40 “An operational, data assimilative model in the Gulf of Mexico: Initial experiences” Lakshmi Kantha
- 11:00 “Coupled physical/biological models” Eileen Hoffman
- 11:30 “Ocean general circulation models and data assimilation” Detlef Stammer
- 12:00 LUNCH
- 1:00 PLENARY DISCUSSION T. Powell  
Suggested topics:  
1) background paper;  
2) previous meeting reports;  
3) emphases for focus groups.
- 2:00 Organization into focus groups  
Suggested foci:  
1) characteristics of the nodes;  
2) computational, technical, and data needs at the central hub;  
3) outreach and educational functions of the hub-node system;  
4) organization and infrastructure within a node-hub structure, including interagency and international partnerships.
- 2:15 Focus groups convene
- 3:00 BREAK
- 3:30 Focus groups reconvene

5:00 ADJOURN  
[Focus group coordinators convene for a brief (!) review of progress].

**Tuesday, May 5<sup>th</sup>**

8:30 PLENARY SESSION. Focus group reports. W. Nowlin

9:30 PLENARY DISCUSSION of focus group reports. T. Powell

10:00 BREAK

10:20 Focus groups reconvene. Participants encouraged to join other groups, if desired.

12:00 LUNCH

1:00 Focus groups reconvene.

2:00 PLENARY DISCUSSION  
Focus group conclusions  
Preliminary recommendations  
The next step(s)

3:00 ADJOURN

## PARTICIPANT LIST

Workshop on Ocean Modeling and Data Assimilation  
4-5 May 1998  
National Science Foundation  
Arlington, Virginia

Robert Armstrong, Princeton University  
Andrew Bennett, Oregon State University  
Rainer Bleck, University of Miami  
Louis Botsford, University of California, Davis  
Piers Chapman, Texas A&M University  
Scott Doney, National Center Atmospheric Research  
Craig Douglas, University of Kentucky  
Christopher Edwards, University of California, Berkeley  
Dale Haidvogel, Rutgers University  
Eileen Hofmann, Old Dominion University  
Lakshmi Kantha, University of Colorado  
Daniel Lynch, Dartmouth University  
James McWilliams, University of California, Los Angeles  
Richard Murnane, Bermuda Biological Station  
Raymond Najjar, Penn State University  
Worth Nowlin, Texas A&M University  
Thomas Powell, University of California, Berkeley  
Warren Prell, Brown University  
Michael Roman, University of Maryland  
Detlef Stammer, Massachusetts Institute of Technology  
Achim Stoessel, Texas A&M University  
Keith Thompson, Dalhousie University  
Leonard Walstad, University of Maryland  
Glen Wheless, Old Dominion University

### **ONR/U.S. Navy:**

Ray Godin, Office of Oceanography of the Navy  
Martha Head, Navy Ocean Office, Stennis  
Jim Price  
Steven Ramberg

### **NASA:**

Antonio Busalacchi, Goddard Space Flight Center  
Lee-Lueng Fu, Jet Propulsion Lab

### **NOAA:**

Ed Harrison, Pacific Marine Environmental Lab  
Ming Ji, National Centers for Environmental Prediction  
T.H. Peng, Atlantic Oceanographic and Meteorological Laboratory  
Steve Piotrowicz, Headquarters  
Carlisle Thacker, Atlantic Oceanographic and Meteorological Laboratory  
Alan Thomas, Headquarters  
Stan Wilson, Headquarters

**National Ocean Partnership Program:**

Cynthia Decker, Consortium for Oceanographic Research and Education  
Rick Spinrad, Consortium for Oceanographic Research and Education

**NSF:**

Jim Ammerman (at NSF then, now returned to Texas A&M University)  
Kendra Daly  
David Garrison  
Don Heinrichs  
Eric Itsweire  
Cliff Jacobs  
Richard Lambert  
Mike Purdy  
Mike Reeve  
Donald Rice

## **APPENDIX B:**

# **A DISTRIBUTED MODELING AND DATA ASSIMILATION CAPABILITY FOR THE OCEAN SCIENCES**

Thomas M. Powell  
Department of Integrative Biology  
University of California, Berkeley  
Berkeley, CA 94720-3140,

Worth D. Nowlin, Jr.  
Department of Oceanography  
Texas A&M University  
College Station, TX 77843-3146,

and

Scott C. Doney  
Climate and Global Dynamics  
National Center for Atmospheric Research  
Boulder, CO 80307-3000

20 April 1998

Prepared for the  
Model and Data Assimilation/Synthesis Workshop  
to be held at the National Science Foundation  
4-5 May 1998

## I. INTRODUCTION

Over the last two years a series of workshops, meetings, informal reports, and one published article (Nowlin 1997) have addressed the proposition that "the present state of" ocean modeling and data assimilation in the United States "is profoundly unsatisfactory".\*\*

What can possibly be the problem? Data is being amassed at an unprecedented rate; satellites and "big" science programs (e.g., WOCE, JGOFS, GLOBEC) are only two examples of many data collection activities. More models exist than ever before in every sub-discipline of the ocean sciences. Computing capability has never been greater. These very characteristics -- large data streams and growing numbers of models -- point to the first broad class of problems: coordination. That is, there is a community-wide need to coordinate the development, testing, maintenance, and sharing of models. The same needs -- for development, testing, maintenance, and sharing -- exist also for large data sets. And if the models are designed to assimilate data, then the need for constructive coordination between modeling and observational communities is even greater. Further, our existing, large, computing capability is deceptive. It masks the estimate that we will need approximately one hundred times our present computing power in a very few years (U.S. WOCE Office, 1997b)! Moreover, coordination endeavors -- very broadly defined -- will have to accommodate the different requirements of each ocean science discipline. For example, circulation and transport models that incorporate data assimilation have been common in physical oceanography for some time (Malanotte-Rizzoli 1996). But coupled biological-physical models are under intense development, and will remain so for a number of years; and only a very few are data assimilative in character (Hofmann 1997, Matear 1995, McGillicuddy et al. 1995a, b). Thus, any approach to remedying our modeling/data assimilation difficulties must certainly include the need for greater computing power, as long as increasing numbers of users with more demanding requirements request computing time. In addition, the users promise to bring an unprecedented diversity of problems from areas of study that, until now, have not heavily consumed computing resources.

In short, arguments to begin a substantial enhancement of modeling and data assimilation capabilities in ALL sub-disciplines of the ocean sciences are compelling. Critical reasons for this need include: existing and new satellites and the data they are (will be) collecting; the massive data sets assembled by WOCE, JGOFS, GLOBEC, etc.; the demands for coordination among modelers, and between modelers and observationalists; and the requirement for greater, yet more diverse, computing capability. These facts are well-documented (U.S. WOCE Office 1997a, b) and are not in dispute.

In September 1997 a small group met informally at the National Science Foundation to consider what might be the next steps to address such community needs. Using recommendations from the meeting reports (U.S. WOCE Office 1997a, b), six notions emerged. First, there are many modeling and data assimilation requirements within the U.S. ocean science community. Some are solely basic research, while others may be operational and applied. Moreover, some research questions can best be addressed by idealized models, while other tasks demand detailed numerical simulations. Any enhancement of our present activities must accommodate this wide spectrum of needs.

Second, the different ocean science sub-disciplines are at different stages of development, as noted above. All can profit from the proposed enhancement in capability, but imbalance

---

\*\* The underscored words in quotes are from the concluding remarks in an influential review by Ursell (1956) on the generation of surface waves by wind. Subsequent investigators noted the "great stimulus to new endeavors" that Ursell's critical review provided (Phillips 1969). Perhaps his words will stimulate another "new" endeavor.

must be avoided. It would be a mistake to focus an inordinate level of resources on the best-developed, or the least-developed, areas of investigation. Many of the most interesting research questions involve linkages between sub-disciplines, so a thoughtful, parallel development seems best.

Third, physical oceanography occupies a critical position in any modeling and data assimilation activity. This is because transport phenomena are so critical to understanding the processes that other sub-disciplines focus upon. Moreover, the technique of data assimilation is much better developed within physical oceanographic model studies. Thus, the other sub-disciplines will look to physical oceanographers for guidance as their own data assimilation questions arise.

Fourth, a related idea, partnerships between investigators from different sub-disciplines will be essential. The coupled physical-biological studies that characterize GLOBEC, or the global CO<sub>2</sub> survey conducted jointly by WOCE and JGOFS, are present-day "big program" examples of such partnerships. Such collaborations must be fostered at all levels of organization, large and small. They must be sought between all the sub-disciplines, because interesting future investigations promise to link disciplines.

Fifth, any proposal to strengthen our present capabilities must provide a structure that facilitates the four previous thoughts -- diversity of modeling and assimilation needs, different stages of development within sub-disciplines, importance of physical oceanography, and requirement for partnerships.

And Sixth, the concrete idea that captured the participants at last September's meeting was that of a "hub-node" system (Nowlin 1997). The "hub" was conceived to have large computational capability that would be used to support scientific teams at the "nodes". Moreover, linkages to data sets, models, and model outputs, as well as the provision of assistance with computational/display/other techniques, likely will be included at the hub. Further, a vigorous visitor/outreach program was envisioned. Some preliminary thoughts on the nature of both the "nodes" and "hub" appear in Nowlin (1997). Several concerns about such a structure were also aired at the September '97 meeting. Perhaps the most serious was the critical need for oversight of "hub" activities to ensure that it remains responsive to the goals of the community. Also, all were uneasy about aspects of the inevitable growth of the "hub" -- especially the pressures for an in-house group of "hub" scientists who might compete with investigators working at the "nodes". These and other elements of any "hub-node" system must continue to be addressed by the ocean science community, should the decision be made to proceed along such lines.

It is our intention here to develop some preliminary descriptions of possible "nodes" -- "strawman nodes". It is hoped that such a description will elicit detailed discussion about the concrete requirements for the "nodes" within the "hub-node" structure. Our description will, by necessity and design, be brief and general. An adequate description would be of NSF-proposal length and detail the models, data, hardware, personnel, etc., that would pertain to a specific group of investigators. A description with that level of detail would not be appropriate here. Such "node" requirements, along with requirements for the "hub" and governance/oversight arrangements, would constitute a preliminary account of the initial stages needed to enhanced modeling and data assimilation capability in the ocean sciences.

## II. NODES

### A. General Considerations

Activities at "nodes" were seen to arise through the normal peer review procedures. A group of investigators with a good idea band together. Synergism would lead to rapid progress that would not be expected if the researchers worked independently of one another. Participants at the September '97 meeting saw "nodes" covering a broad spectrum of modeling and data synthesis activities in the ocean sciences. Some of the suggested nodes were:

- population dynamics modeling
- ecosystem modeling
- paleoceanography modeling
- surface layer modeling
- nested model development
- access to, and use of, real-time coastal observations
- seasonal-decadal climate modeling
- geochemical models
- carbon dioxide models
- productivity estimation/bio optical modeling
- global ocean processes/circulation
- long-term climate modeling
- coupled ice-ocean modeling
- numerical techniques/signal processing

"Nodes" would have limited durations...say three to five years, with the possibility for renewal (though continuation for indefinite periods drew negative responses). The "nodes" would be expected to have more tightly defined emphases than the present big programs (WOCE, JGOFS, GLOBEC). Some of the "nodes" might be very closely proscribed [development of sigma-, or s-coordinate models; unstructured grids -- see U.S. WOCE Office (1997b) for some further examples]. And some of the "nodes" might have broader missions -- development of coupled biological-physical ecosystem models. The "nodes" were seen as programs of intermediate scale -- mesoscale programs. They would involve fewer investigators than, say WOCE or JGOFS, with less administrative overhead. But they would be substantially larger than a single-investigator project. We have persisted in placing "node" within quotation marks, because the word carries the implication of a single site. This "one-place" idea seems too narrow; there was strong support for the notion of distributed nodes, with investigators at a number of locations. Modeling and data assimilation appear to be activities that can be profitably pursued in this distributed fashion; and advances in software and hardware seem likely to make such activities even easier in the future.

As nodes develop, thought will have to be directed to the overall organization of this structure. All the nodes will face similar "outward-looking" and "inward-looking" problems: outward to the larger community, and inward within the node and to the "hub". Is the changing mix of nodes achieving larger community objectives? Do the users outside the immediate node-group have easy access to model advances and data sets? Should outside users be accommodated with workshops, hands-on tutorials, etc., directed toward the use of node products? It seems likely that an advisory committee will be formed to oversee the scientific directions of the "hub-node" structure. One of the functions of such a committee might be to see that links to the "hub" are functioning smoothly and constructively. Just like a webmaster, will we need a "hubmeister"? Will each node require its own "hubmeister"? Are appropriate connections being formed between the

"hub-node" structure and the modeling and assimilation activities in other agencies? These general considerations, and more, will occupy the attention of the community as we move toward enhancing our modeling and data assimilation capabilities.

We now turn to specific examples of activities at nodes. They are, of course, speculative, and hypothetical. And contrived to meet the requirements advanced in the Introduction. Most of all, they are advanced to begin discussion of the capabilities the ocean science community will demand in the future.

## **B. The Coastal Ocean**

A recent review (Haidvogel and Beckmann 1998) noted that one must require numerical models in the coastal ocean to be flexible and highly optimized. Flexible, because so many physical phenomena are found in the coastal ocean -- fronts of all kinds, currents that depend strongly on vertical and onshore-offshore position, coastal trapped waves, strong vertical stratification, intense generation of turbulence and mixing at bottom and surface boundary layers, and many more. Optimized (for particular regions and processes) because dependence on specific details of irregular coastlines and variable bathymetry is so great. Moreover, because of the social and economic importance of the coastal zone, and proximity to such a large fraction of the world's population, the ready ability of models to address many applied problems (e.g., oil spills, pollution episodes, fisheries management and storm damage) will be another distinguishing feature of successful models.

A Coastal Ocean Modeling Node might consist of several (say four to six) distinct groups. One group might devote its attention to GENERAL PROCESSES AND ALGORITHMS. A second group might have overall responsibility for LINKAGES. Moreover, there might be one "linkage" sub-group addressing PHYSICAL PROCESSES, like the incorporation of mixed-layer models (Large et al. 1992, Mellor and Yamada 1982), or nesting high-resolution nearshore formulations into general process models. In addition, there might be one or more "linkage" sub-groups devoted to INTERDISCIPLINARY COUPLING; for example, incorporating the special characteristics of sea-ice, or sediment transport models, into the general process formulations. (Below, we give special attention to another interdisciplinary activity -- coupling of physical and biological processes.) There might also be a TEST PROBLEMS group (sub-group?) that would suggest a series of test calculations against which various models, formulations, etc., could evaluate themselves. It might be valuable to keep the general process and algorithm group separate from the test problem group -- friendly competition could spur more rapid development. Finally, a DATA AND ASSIMILATION GROUP, to include observational oceanographers, should have close connections to all of the other allied groups and sub-groups. Moreover, the task of constructing the assimilative formulations (the "backward" models) associated with the models constructed by the other groups (the "forward" models) would be a shared responsibility. The data assimilation investigators and scientists from the other groups would be expected to develop a modeling and assimilation scheme jointly. The data and assimilation team also would be expected to develop and maintain links to data archives.

## **C. Coupled Physical-Biological Models**

A node might be devoted to the development of coupled Physical-Biological Models and the construction of schemes to assimilate biological (as well as physical) data into such models. The node participants would be guided by the observation of Hofmann and Lascara (1998) that the difficulty in this activity arises when one attempts "... to combine circulation and biological processes in a single model". These authors further observe that the requirements for temporal and spatial resolution for one activity (describing biological

processes) may differ from the other (modeling physical phenomena). In fact, the requirements may be contradictory! Linkage, then, is the dominant focus.

The node could consist of three primary groups. The first might address PHYSICS/CIRCULATION/TRANSPORT. A critical aspect of this group's efforts would be its connection to other physical oceanographic activities. Because many of the most pressing biological questions arise in coastal areas one could imagine this group could be connected to those developing coastal circulation models. One strategy would be for the group to conceive of themselves as "users" -- educated, savvy users, but not developers. The users might employ a basic circulation model from other researchers, modifying and adding components as necessary. This group might have a large number of members. Indeed, four foci are evident from problems facing investigators in this emerging field. First, embedding mixed layer models in circulation models will be a critical task for this group. Biological activity is highest in near-surface water and vertical transport acting over short time scales contributes greatly to variability in biological production. Thus, biological simulations will need accurate representation of mixing processes and net vertical motion in the mixed layer. Second, mesoscale features are pervasive and energetic in the sea; and they have demonstrable biological effects (McGillicuddy et al. 1995a, b). So the physical process group must emphasize nesting of high-resolution, eddy-resolving transport models. Third, many compelling biological models are Lagrangian in character, notably individual-based models, but most physical process models are Eulerian. So the physics group will face the task of transforming results seamlessly and flexibly between the two pictures. Fourth, and perhaps most important, the physical models must be readily accessible to biological partners. This means that "hooks" must be built into the physical code upon which biological process formulations can be hung.

The second group might be primarily responsible for the formulation of models for basic BIOLOGICAL PROCESSES. This endeavor is developing rapidly, but a broad spectrum of models will be necessary into the foreseeable future because a consensus about which are the "best" has yet to emerge. Some attractive models are bound to be simple and idealized, while others may be detailed and complex. One can imagine a continuing emphasis on many levels of organization, from the individual level to the highly aggregated biomass spectrum models. Further, it seems likely that consideration of various life-history stages for planktonic organisms will increasingly be necessary. Moreover, efforts must continue on many trophic levels: phytoplankton, microbes, other invertebrate carnivores, fish, and, perhaps seabirds and marine mammals. Organization around foci (like the physical process group) is more difficult for the biological partners because the area of biological modeling is developing so rapidly. But a successful biological-physical modeling node must maintain tight links between "the physics" and "the biology". Also, the need for outreach to the larger biological oceanography community that is less exposed to models than that of the physical oceanographers is particularly acute. A continuing series of hands-on tutorials in the use of models would be most welcome -- and should be the responsibility of the biological process group.

The third group will have responsibility for DATA AND ASSIMILATION. This group will have all the characteristics noted for the data and assimilation group in the previous coastal ocean node -- containing observationalists, close connections to the other allied groups, shared responsibility for the assimilation schemes, and links to data archives. In addition, certain data types -- especially biological data -- are "new" insofar as they have not been used in model-data comparisons of any kind -- let alone in assimilative schemes. Such data include acoustic information (with very high data rates) and images from in-situ plankton sensing devices (Video Plankton Recorder -- Gallagher et al. 1996). These new data show great promise and the data and assimilation group should anticipate their use in future modeling efforts. Finally, the assimilation task for biological data (into coupled

physical-biological models) can be expected to provide a surprise or two -- even to the experienced practitioner. Biological data has inherently different characteristic time and space scales from those seen in physical data; and non-linearities found in biological process descriptions differ from those in physical equations. Though the problem of data assimilation in this general case is no different in principle from that for "the physics", the prudent investigators in the data and assimilation group will be watchful for new and challenging difficulties.

#### **D. Marine Geochemistry and Biogeochemistry**

The fields of marine geochemistry and biogeochemistry are concerned broadly with the chemical transformation and transport of material within the ocean as well as the mass fluxes across the sedimentary and air-sea boundaries. Unlike, for example, the atmosphere, a large fraction of the chemical processes in the ocean are biologically mediated, and the chemical field both contributes to and relies on advances in the related areas of ocean physics, biology and geology. The last decade has seen a dramatic increase in the quantity of high quality chemical data (e.g. inorganic carbon from JGOFS-WOCE global survey) and the range of properties that can be measured (e.g. characterization of bulk dissolved organic carbon composition; trace metal concentrations and speciation). This wealth of data will be further compounded with the current advent of a variety of automated chemical analysis techniques for underway, towed, moored, and autonomous platforms. A major challenge confronting the field is to develop the appropriate numerical tools to interpret these often exquisite and painstaking analytical measurements within an overall oceanographic context.

There are a number of ways to divide the computationally intensive data analysis and modeling aspects of chemical oceanography, and we have chosen a time-space structure for the proposed nodes, realizing that there maybe considerable overlap for some processes. Differing from the coastal and biological proposals, physical modeling and data assimilation would not have separate nodes but would be required sub-components of three (and perhaps all) of the four proposed nodes for marine geochemistry and biogeochemistry.

Global Biogeochemical Cycles Node. The central focus would be on quantifying the factors controlling the large-scale distributions and fluxes of nutrients, oxygen, dissolved and particulate inorganic and organic carbon. A primary goal would be to better constrain the governing mechanisms of surface export fluxes, subsurface remineralization and sediment processes. This node would be dependent on the large-scale data sets from JGOFS and WOCE along with improved global physical circulation models (link to an OGCM node? see D. below), but would also help constrain physical flow fields, particularly if water-mass tracers such as tritium, radiocarbon and chlorofluorocarbons are incorporated. Sub-group foci would include applications to paleoceanographic model reconstruction (e.g. glacial or interglacial), climate variability and anthropogenic change, including potential feedbacks on atmospheric CO<sub>2</sub>.

Ocean Margins and Terrestrial-Marine Interactions Node. Coastal regions play a key role in a number of important marine biogeochemical processes (e.g. denitrification, carbon burial, authigenic mineral formation). The integrated global impact is often poorly characterized because of regional variability, as is the mass exchange between the coastal and open oceans. This is one area which could benefit greatly from embedded regional models (link to the coastal ocean node, A., above). Considerable effort is also required to improve estimates of the natural and anthropogenic (e.g. eutrophication, pollution) chemical inputs to the coastal zone; areas of required focus include atmospheric deposition,

riverine discharge and estuarine transformations, advective flows through margin sediments and coastal aquifers, and sedimentary processes.

Upper-Ocean and Marine/Atmospheric Boundary Layer Node. The seasonal boundary layer of the ocean is of intrinsic interest as the primary zone of biological production and the interface to the troposphere. Modeling difficulties arise, however, because of the rapid chemical and biogeochemical rates and short space and time scales associated with boundary layer mixing and mesoscale circulation. The potential data stream for this region from satellite remote sensing and automated chemical analyzers is also quite large. An important set of scientific questions centers on characterizing the resulting variability and the rectification into the mean state. Better dynamically based models are also needed for: photochemical processes; nitrogen fixation; formation, remineralization and modification of organic matter (dissolved and particulate); relationship among trace metals and micronutrients, community structure and export flux; and the cycling and air-sea exchange of trace gas species and aerosols (e.g. DMS, organohalides).

Process Models Node. A number of oceanographically relevant modeling topics, mostly process oriented studies, must not be neglected. The potential members of a process node are rather disjoint but have many computational and/or data assimilation requirements in common with the nodes discussed above. Some topics are still based in fluid dynamics but could fall in this group because they are very small scale, such as direct numerical simulations (DNS) (e.g., surface wave breaking; plankton-plankton interactions), or large-eddy simulations (e.g., Langmuir and boundary layer circulation). Other areas are less related but still important, for example thermodynamic and flow models of water-rock and water-sediment interactions along mid-ocean ridges and flanks, or chemical models relating dissolved organic matter structure to chemical reactivity.

## **E. General Circulation of the Ocean.**

The general circulation of the ocean is an essential element of the oceanographer's view of the sea: it is three-dimensional and planetary in scale...and it is grossly undersampled, especially below the surface layers. Large-scale data collection efforts (e.g., WOCE, JGOFS, TOGA, satellite missions) have provided a vast, continuing store of information. And our ability to construct adequate models of the ocean general circulation increases each year. We have not, however, performed a satisfactory synthesis of this information. The general circulation is the quintessential ocean modeling and data assimilation problem. And there is every reason to believe that we possess the technical skill to perform this synthesis.

An Ocean General Circulation Model (OGCM) Node might have as many as six groups. First, a group devoted to MODEL CONSTRUCTION AND DEVELOPMENT (the "forward" model) would be complemented by a second group addressing DATA ASSIMILATION TECHNIQUES (the "backward" model). A third group emphasizing REGIONAL MODELING foci might have several sub-groups -- a tropics/equatorial division, a mid-latitude/subtropics team, and a high latitude (especially Southern Ocean) branch. Fourth, development of several MULTI-DISCIPLINARY LINKAGES would be especially valuable. For example, in Section D, above, we noted the critical role that the ocean general circulation plays in global biogeochemical cycles. And coastal ocean workers who tackle the nesting problem (high-resolution, nearshore models embedded in lower resolution, larger-scale models) will profit from reliable OGCM output when considering how information should be exchanged between both scales. Fifth, a group devoted to COMPUTATIONAL INNOVATIONS seems necessary because the continuing assimilation of enormous data sets will surely tax our present computational resources (see Semtner's estimate in U.S. WOCE Office 1997b). The use of large numbers of relatively inexpensive computing units connected in parallel may be the avenue that this activity must

exploit. Finally, and perhaps most importantly, close connections must be forged between the modeling teams and collaborating OBSERVATIONAL groups. Such collaborations might include: satellite/remote sensing investigators; chemical tracer experts; and teams concerned with floats and drifters of various kinds.

### III. CONCLUDING REMARKS

We have devoted most of our attention here to examples of possible nodes. Properly organized and coordinated, a changing mix of such nodes can effectively address the modeling and data assimilation requirements set out previously (U.S. WOCE 1997a,b; Nowlin 1997), and earlier in this paper. Furthermore, such a structure possesses sufficient flexibility to satisfy the important demands for diversity of approaches and techniques, as discussed in the Introduction. It would also allow partnerships to develop at the right level -- between scientists; and for the right reasons -- scientific ones. Of course, there are many more possible, important nodes beyond the four listed. In particular, the report on global circulation modeling (U.S. WOCE Office 1997b) noted a number of foci [ e.g., fixed grid models (z-, or sigma- coordinates); and unstructured horizontal grids] that are more specific than those suggested here. But a "hub-node" structure would be sufficiently flexible to accommodate tightly-defined nodes, too. Alternatively, such specific foci could function as sub-groups within larger nodes that have broader missions.

We have given little attention to questions of scientific organization, management, and leadership. These will ultimately prove to be very important. A recent review of NSF's Science and Technology Centers (STCs) strongly emphasized that "Scientific leadership is the key to a successful center, especially if it is managing highly complex, multi-investigator, multi-institutional research" (NAS 1996). Since some of the envisioned nodes might have sizes, even lifetimes, of the small-to-medium STCs, this caution should apply to the nodes and, certainly, to the overall "hub-node" structure. We also note that the ocean science community does have a history of handling large, distributed programs/facilities with long lifetimes. Examples include the Ocean Drilling Program, the UNOLS system, and large programs like WOCE, JGOFS, and GLOBEC. Nonetheless, beyond calling for a Scientific Advisory Committee, it may be premature to spend much time here on questions of organization, management, and leadership of the nodes. But, even at this early stage, one should recommend that peer review for the initial nodes must carefully consider leadership, management, etc., of the proposed activity.

#### A. The "Hub"

The concluding section of Nowlin's (1997) article notes several possible characteristics of the "hub": large computing capability; the ability to handle data in near real time, as well as the enormous data sets that have been collected recently (and historically); access to model code and output, coupled model/data products, and experts who can aid the user with technical tasks; and a vigorous outreach/visitor program that will attract many in the ocean science community. The first of these (large computing capability) is the most crucial building block. The other characteristics become important when a central computing facility -- more powerful than any single user, or group of users (i.e., a node) can justify -- is a reality. Frankly, the entire program is scarcely worth pursuing unless an effective central computing "hub" emerges to support the scientific activities at the nodes. Semtner's estimate (U.S., WOCE Office 1997b) defines the basis for "effective" -- a factor of 100 relative to the present. That is, a rise from 10 gigaflops (the approximate present usage) to roughly 1000 gigaflops must occur in a very few years. Once the commitment to a major computing facility has been agreed upon, then the other characteristics, above, would be essential to the effective and responsive functioning of a "hub" within a "hub-node" structure. In the Introduction we mentioned concerns about a large "hub" --

responsiveness to community goals and competition between community users and "hub" scientists. Again, it seems premature to suggest solutions to these potential difficulties, though the leadership of a "hub-node" structure would certainly have to face them early on.

## **B. Timing**

All considerations point to the need for a rapid initiation of the elements of this project. The problems that demand a greatly enhanced ocean modeling and data assimilation capability will not go away simply by our neglect of them. We will still have to cope with vast (and growing) data sets, assimilating them into our models. We will still have to accommodate a rapidly growing community of modelers (including their proposed assimilation activities) from all the sub-disciplines within the ocean sciences, few of whom have demanded substantial computing support in the past. And the clamor will grow quickly for greatly increased computing resources within our discipline.

The "hub-node" structure does present a "ramp-up" strategy toward implementation. Perhaps one could initiate the project with an Announcement of Opportunity for, say, two or three nodes. If the selection criteria emphasized the importance of diversity, then two or three very different node projects might emerge (at least one in a rather highly-developed modeling/assimilation posture). This would allow a Science Advisory Committee to incorporate the expertise and experience of node-users into the criteria for the selection of the "hub" -- the most critical single element in the "hub-node" structure, certainly the most expensive, the one with the longest lifetime (and demand for the greatest commitment of long-term support, presumably via block-funding), and the obvious next step. Once the "hub" is "on the tracks", further calls for nodes could be released until a full complement is attained. This process might take a decade.

Releasing an initial call for nodes is the right first step because it will signal to the community that this is "node-driven science". The questions that motivate the users of the "hub", will propel the overall program. Though the most constructive scientific discourse always involves give-and-take -- in this case between "hub" and nodes -- requirements from the nodes should shape the "hub", and not the converse.

We argue that this first step, a call for node projects, should be taken soon.

## **ACKNOWLEDGEMENTS.**

We are grateful to Dale Haidvogel, Eileen Hofmann, Detlef Stammer, and Carl Wunsch for material from which sections of this report were derived. We thank Maureen Reap and Kristin Bishop for preparation of the manuscript.

---

## **REFERENCES.**

Gallagher, S.M., C.S. Davis, A.W. Epstein, A. Solow, and R.C. Beardsley. 1996. High-resolution observations of plankton spatial distributions correlated with hydrography in the Great South Channel, Georges Bank. *Deep-Sea Res., Part II.* 43: 1627-1663.

Haidvogel, D.B., and A. Beckmann. 1998. Numerical models of the coastal ocean. In K. Brink and A.R. Robinson (eds.). *The Sea*. Vol. 10, J. Wiley and Sons, New York.

- Hofmann, E.E., and C.M. Lascara. 1998. An overview of interdisciplinary modeling for marine ecosystems. In K. Brink and A.R. Robinson (eds.). *The Sea*. Vol. 10, J. Wiley and Sons, New York.
- Large, W.G., J.C. McWilliams, and S.C. Doney. 1992. Oceanic vertical mixing: a review and a model with a non-local boundary layer parameterization. *Rev. Geophys.* 32: 363-403.
- Malanotte-Rizzoli, P. 1996. *Modern approaches to data assimilation in ocean modeling*. Elsevier, Amsterdam.
- Matear, R.J. 1995. Parameter optimization and analysis of ecosystem models using simulated annealing. A case study at Station P. *J. Mar. Res.* 53: 571-607.
- Mellor, G.L., and T. Yamada. 1982. Development of a turbulent closure model for geophysical fluid problems. *Rev. Geophys.* 20: 851-875.
- McGillicuddy, D.J. Jr., J.J. McCarthy, and A.R. Robinson 1995a. Coupled physical and biological modeling of the spring bloom in the North Atlantic (I): Model formulation and one-dimensional bloom processes. *Deep-Sea Res.* 42: 1313-1357.
- McGillicuddy, D.J. Jr., A.R. Robinson, J.J. McCarthy. 1995b. Coupled physical and biological modeling of the spring bloom in the North Atlantic (II): Three-dimensional bloom and post-bloom processes. *Deep-Sea Res.* 42: 1359-1398.
- NAS. 1996. *An assessment of the National Science Foundation's Science and Technology Centers*. Committee on Science, Engineering, and Public Policy. National Academy Press, Wash., D.C.
- Nowlin, W.D., Jr. 1997. U.S. ocean science needs for modeling and data synthesis: status of a community assessment. *Oceanography* 10:135-140.
- Phillips, O.M. 1969. *The dynamics of the upper ocean*. Cambridge Univ. Press.
- Ursell, F. 1956. Wave generation by wind. pp. 216-249 in Batchelor, G.K., and R.M. Davies (eds.). *Surveys in mechanics*. Cambridge Univ. Press, New York.
- U.S WOCE Office. 1997a. Assessing ocean modeling and data assimilation requirements. 1. Report on a workshop to discuss needs for an ocean data assimilation center. May 1997.
- U.S WOCE Office 1997b. Assessing ocean modeling and data assimilation requirements. 2. Report on a workshop to discuss community needs for ocean global circulation modeling. May 1997.